RESEARCH ARTICLE

# Neural networks and differential evolution algorithm applied for modelling the depollution process of some gaseous streams

**Silvia Curteanu · Gabriel Dan Suditu ·
Adela Marina Buburuzan · Elena Niculina Dragoi**

**Abstract** The depollution of some gaseous streams containing *n*-hexane is studied by adsorption in a fixed bed column, under dynamic conditions, using granular activated carbon and two types of non-functionalized hypercross-linked polymeric resins. In order to model the process, a new neuro-evolutionary approach is proposed. It is a combination of a modified differential evolution (DE) with neural networks (NNs) and two local search algorithms, the global and local optimizers, working together to determine the optimal NN model. The main elements that characterize the applied variant of DE consist in using an opposition-based learning initialization, a simple self-adaptive procedure for the control parameters, and a modified mutation principle based on the fitness function as a criterion for reorganization. The results obtained prove that the proposed algorithm is able to determine a good model of the considered process, its performance being better than those of an available phenomenological model.

**Keywords** Differential evolution · Neural network · Local search · Backpropagation · VOC adsorption · Hypercross-linked polymeric adsorbents

## Introduction

It is well known that volatile organic compounds (VOC) are major contributors to the formation of the photochemical ozone and secondary organic aerosol, which would result in serious health and environmental problems. In this sense,

S. Curteanu · G. D. Suditu · A. M. Buburuzan · E. N. Dragoi (✉)
Faculty of Chemical Engineering and Environmental Protection,
"Gheorghe Asachi" Technical University of Iasi, Bd. Prof. dr. doc.
DimitrieMangeron, No. 73, 700050 Iasi, Romania
e-mail: elenan.dragoi@gmail.com

national environmental protection agencies are forcing the chemical industry to reduce the emission of these compounds both in wastewater and airstreams and, thus, emissions of VOCs are becoming one of the most stringent environmental challenges in many industrial processes (Odabasi et al. 2005; Hu et al. 2009; Silvestre-Albero et al. 2010).

In general, VOC emissions include a wide range of chemical substances such as aromatics, various chlorinated hydrocarbons, aldehydes and ketones, alcohols, organic acids, esters, and aliphatics. Various methods are used to eliminate volatile organic compounds, the most efficient by an economic point of view being those that involve separation and reuse of VOCs, such as absorption, adsorption, condensation, and membrane separation (Matros et al. 1988).

Elimination techniques of VOCs can be divided into two categories: (1) destructive techniques (thermal and catalytic oxidation, photochemical oxidation, and biodegradation), which destroy the undesirable compounds, and (2) recuperative techniques (adsorption, absorption, condensation, and membrane separation), allowing their recovery (Khan & Ghoshal 2000). In general, the methods other than adsorption (especially condensation) are effective when VOC concentrations are relatively at high levels (>1 %). Therefore, adsorption has been found to be effective, with higher selectivity and relatively higher capacity for VOC, even at low concentration levels (ppm or sub-ppm) (Gupta & Verma 2002; Kim et al. 2006; Lillo-Rodenas et al. 2006; Shim et al. 2006).

The emission of solvent vapors from the industrial processes can cause not only severe air pollution but also great loss of valuable chemicals. Therefore, proper recovery of volatile solvent vapors from industries will help in reducing production costs, saving energy, and protecting the environment (Khan & Ghoshal 2000; Shim et al. 2006).

Related to the adsorption process as one of the most applied for VOC removal, its complexity derives from the fact that it is influenced by the type and concentration of VOCs in the

initial pollutant flow, gas flow rate, relative humidity of gas stream, nature of the sorbent and its physicochemical characteristics, length of adsorbent bed in the column, residence time, and temperature of the adsorption process. On the other hand, the popularity of the method arises from the possibility to obtain and use different materials.

Taking into account that determination and characterization of the conditions for VOC removal from waste gas streams are laborious, a cheap and easy to apply alternative is represented by the simulation of process removal. In this way, it is possible to replace or, at least, to plan experiments that are material, energy, and time consuming. In addition, predictions can be made in real time in situations where accidental pollutions appear with these substances and useful information can be obtained about the analyzed processes. Also, a good model is a prerequisite in optimal control procedures.

Artificial neural networks (NNs) represent efficient tools for chemical process modelling, being recommended when classical approaches fail to provide acceptable models (results). In this context, NNs are viable modelling alternatives, since they have the property of universal approximation (Subudhi & Jena 2009). A NN is a computational structure inspired by the biological brain, which consists of a set of interconnected processing elements called neurons (Xin 1999). It has a series of advantages which include, among others, modelling non-linear relationships, flexibility, parallel processing, learning and self-learning, and fault tolerance (Hernandez et al. 1998; Patan 2008). Besides other artificial intelligence methods (Caicedo et al. 2013) NNs can be applied to solve specific problems (Vassileiou et al. 2012), their advantages making them suitable for different types of application. At the same time, they are good alternatives to different solutions such as those found in environmental decision (Matthies et al. 2007) or system stabilization (Precup et al. 2007).

Some examples of using neural networks for processes involving VOC adsorption or their detection can be found in literature. Tchoupo and Guiseppi-Elie (2005) used NNs to simulate and classify the adsorption of four gases (toluene, benzene, heptane, and butanol) on six polymer sensors. The design of new materials for sensors was made by Han et al. (2005), who reported the findings of a research on nanostructured materials, sensitive for the detection of volatile organic compounds and nitro-aromatics, the analysis being performed based on pattern recognition using NNs and principal component analysis. An electronic system for detection and identification of VOCs was designed and developed by Srivastava (2003) through a series of gas sensors based on tin oxide and using NNs. Backpropagation neural networks were used to identify relevant VOCs for environmental monitoring such as 2-propanol, methanol, acetone, methyl ethyl ketone, hexane, benzene, and xylene. Other applications of NNs in the field of chemistry and chemical engineering can be found in the review of Pirdashti et al. (2013).

In the current work, the VOC adsorption process on activated carbon and polymeric materials is approached, aiming to retain pollutants from gas streams with high concentrations of VOCs. Adsorption on granular activated carbon is one of the well-known and efficient methods used for VOC recovery from a gaseous stream. However, activated carbon has some important drawbacks such as the self-firing of the adsorbent in the bed, pore blocking, and difficulty in the regeneration of the spent adsorbents. Therefore, it is necessary to develop new adsorbent materials to separate and recover VOCs from polluted airstreams. For this purpose, in the previous works, the potential of two non-functionalized hypercross-linked polymers type Macronet, MN 202 and MN 250, was studied for the adsorption of $n$-hexane vapors (Halteta Buburuzan et al. 2009; Buburuzan et al. 2009; Buburuzan et al. 2010). Using these materials along with charcoal, a series of experiments were performed in different reaction conditions for determining the efficiency of the process.

In this article, the role of the experiment was to create a complete database from which a good neural network model can be derived. The problem of determining the optimal architecture of the neural network is also approached, the proposed solution making use of a relatively new evolutionary algorithm (EA) named differential evolution (DE).

The simplicity of NNs is deceptive because training is a difficult task when an optimized architecture is desired. In addition, determining an optimal topology is problem dependent, various contradictory rules being encountered in literature. In order to solve this aspect, the NNs are combined with EA, and the advantages of this approach are as follows: (1) simultaneous evolution of several defined features, (2) a flexible definition of performance criterion, and (3) the possibility of coupling an EA with a learning algorithm (Floreano et al. 2008).

DE is a bio-inspired technique capable of working with a variety of objective functions including non-differentiable, non-linear, or multimodal functions (Pant et al. 2009). It is a stochastic, population based, powerful, and easy to implement optimization algorithm which was successfully applied to solve different problems from various domains such as engineering design, control, scheduling, chemical engineering, decision making, or image processing (Feoktistov 2006). As examples of EA-NN combinations, one can mention our previous contributions in determining optimal neural networks using the improved non-dominant sorting genetic algorithm (Furtuna et al. 2012; Llanos et al. 2013) or different variants of DE (Dragoi et al. 2011, 2012a).

In the current paper, a hybridization of the self-adaptive DE algorithm is proposed and tested for the neural modelling of a depollution process of some gaseous streams containing volatile organic compounds. At the end of each generation, the best individual found is improved by randomly applying the random optimization or backpropagation (BK) algorithm.

This new hybrid DE algorithm (called hybrid self-adaptive differential evolution with neural network (hSADE-NN)) is used for determining the optimal neural network which can model the considered process. Consequently, the NN represents the problem being solved by the optimizer, a simultaneous structural and parametric optimization being performed.

The main contributions of this paper are related to the creation of a consistent dataset obtained by adsorption of hexane into different categories of materials and neural network modelling performed using an optimization procedure based on a new variant of a hybrid differential evolution algorithm. A comparison is also made with one of our own previous DE versions and with a kinetic model, proving the efficiency of hSADE-NN and its adequacy for the approached process and, also, justifying the creation of this new version.

## Methods

The volatile organic compound used in this work was *n*-hexane, with an analytic grade of 99.9 %, purchased from the Chemical Company and used without further purification. Dynamic adsorption of *n*-hexane vapors was studied comparatively on two categories of adsorbents: the granular activated carbon AC 20 which has a bituminous origin and two types of non-functionalized hypercross-linked polymeric resins Hypersol-Macronet TM, MN 202 and MN 250. Both polymers consist of macroporous polystyrene crosslinked with divinylbenzene without functional groups.

The principle of Hypersol-Macronet resins is that the porosity is not introduced during polymerization but during a post polymerization process in which the polymer is hypercross-linked in a swollen state and results in two special properties: (1) a very high surface area, comparable with that of activated carbon, and (2) swelling by liquids and gases that do not normally solvate polymeric matrices.

The polymers from class MN 200 tend to be hydrophobic and may be considered for the removal of organic pollutants. All adsorbents were supplied by Purolite International Limited country offices of Purolite Corporation in Bucharest, Romania (Buburuzan et al. 2010).

Vapor breakthrough experiments of *n*-hexane were carried out in a 1.0-cm-diameter quartz column, using different bed heights of adsorbents, concentrations of *n*-hexane-air mixture, and temperatures at a constant flow rate. The temperature of the adsorbent bed was controlled and measured by a thermocouple placed as close to the absorber column as possible. The concentration of *n*-hexane vapor in the airstream, before and after the adsorption column, was measured by using a gas chromatograph (GC) Carlo Erba 4200 Series. The GC is equipped with a six-way valve, a column with 10 % DC-200 Chromosorb PNAW, a flame ionization detector (FID), and an integrator. The injector, oven, and detector temperature were maintained at 170, 110, and 170 °C, respectively. Hydrogen gas was used as fuel and nitrogen as carrier gas. The calibration curve was prepared by injecting known amounts of VOCs into a sealed bottle equipped with a Teflon septum, according to the standard procedure.

To study the effect of concentration on the breakthrough time and adsorption capacity, the concentration of *n*-hexane vapors in the gaseous stream was varied at 7, 11, and 16 mg/L at the flow rate of 130 mL/min. The bed length of adsorbents in the column was 1 cm for all the three adsorbents studied.

To find out the effect of bed length on the breakthrough and adsorption capacity of *n*-hexane vapors, the experiments were carried out in 20 mg/L of *n*-hexane passed at the flow rate of 130 mL/min at different bed lengths, i.e., 1, 1.5, and 2 cm of MN 202, MN 250, and AC 20, respectively. To examine the effect of temperature on the *n*-hexane adsorption process, the temperature was varied at 30, 40, and 50 °C, with 11 mg/L of *n*-hexane, a flow rate of 130 mL/min, and 1 cm bed length for all the three adsorbents.

This kinetic study of VOC adsorption on different materials leads to a dataset containing 675 data, which was prepared for modelling purpose. Neural network modelling is an appropriate tool for this process because important consumption of time is necessary for obtaining optimal retention values when high quantities of VOC are present in gaseous effluents. Instead of a significant number of experiments, the resulting model can be used to predict the optimal conditions for depollution.

For NN models, contact time, initial concentration of VOC, specific surface of the adsorbent, thickness of the adsorption material, and temperature were chosen as inputs and elimination efficiency as output. The analyzed parameters and their limits of variation were the following:

- Type of adsorbent, quantified through surface area BET. Taking into account the nature of this parameter, the variation interval only includes three values, namely 669.4 for AC 20, 789.38 for MN 202, and 980.26 for MN 250.
- Contact time, 0÷220 min.
- Bed length, 1, 1.5, and 2 cm.
- Temperature, 25, 30, 40, and 50 °C.
- Initial concentration of VOC in the gaseous effluent, 7, 11, and 16 mg/L.

In this work, the neural network acts as a model for the considered process, while the modified DE is an optimizer with the role of determining the best possible model. The main process parameters are considered as inputs and outputs, respectively, while the experimental data are used for training and testing the performance of the generated models. A simplified schema of this workflow is presented in Fig. 1, where
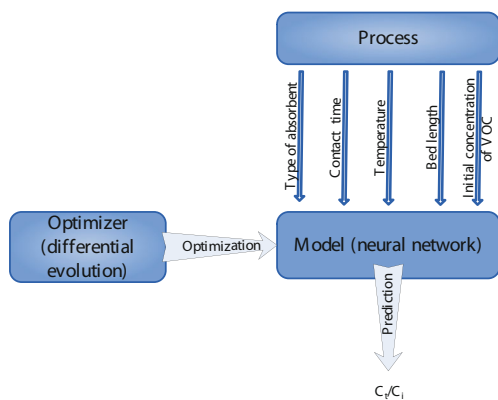
**Fig. 1** Simplified schema of the modelling and optimization workflow

$C_t$ is the outlet concentration at arbitrary time and $C_i$ is the inlet concentration.

After the data were gathered, a min-max normalization procedure was applied. This is one of the most used methods of data pre-processing, which leads to satisfactory results since it can reduce the estimation error and time calculation and can improve the capability of discriminating high-risk software (Leeghim et al. 2008). The linear interpolation formula used to normalize the data is described by Eq. 1:

$$x_{norm} = min_t + (max_t - min_t)\frac{x - min}{max - min} \qquad (1)$$

where $x_{norm}$ represents the normalized value of $x$ parameter, $min_t = -0.9$ and $max_t = 0.9$ are the values determining the target interval, and min and max describe the interval in which $x$ takes values.

## Developing an optimal neural network model with a hybrid version of differential evolution algorithm (hSADE-NN)

Since the most simple, easy to use neural network with the function of universal approximator is represented by the feed forward multilayer perceptron (MLP), this type of network was chosen for the current study.

Due to a series of aspects such as a high number of parameters, various possible combinations, multiple training algorithms (each having various drawbacks), and lack of consistent rules, determining the optimal topology and internal parameters of neural networks is a difficult task (Curteanu & Cartwright 2012). In the current paper, a neuro-evolutionary approach based on a new version of DE algorithm is used to solve this problem. All the alterations performed to the DE, along with the optimizer-model combination proposed in the current work, which concretized into a new algorithm, are reflected into its name: hybrid self-

adaptive differential evolution with neural network (hSADE-NN). The main characteristics of hSADE-NN, the basic ideas, and the motivation for specific choices of some parameters are further presented.

Various studies successfully employ DE for training neural networks, the simultaneous optimization of topology and training being scarcely encountered in literature (Subudhi & Jena 2009; Chandra & Yao 2006; Dragoi et al. 2012b, 2013). This can be explained by the fact that optimization requires a large search space and a high number of parameters which, in their turn, have an impact on the performance of DE, even if this algorithm can be successfully used when the network is very large, when the problem has many local minima, or where there is much noise in the data (Zarth & Ludermir 2009).

Although very powerful and with several attractive features, it has been observed that in some cases DE is inefficient. It inherits the crucial flaws of the evolutionary computation by having a slow convergence rate and low accuracy, especially in noisy environments or when the solution space is hard to explore (Peng & Wang 2010; Neri & Tirronen 2010).

In order to improve the DE algorithm and eliminate or reduce the effect of the inherited problems, a series of modifications were proposed by various researchers. These modifications correspond to the main directions of research related to (1) basic DE research (control variables, perturbation and diversity enhancement, controlling the vector population), (2) research specific to the problem domain (where the problem formulation and the manner in which DE can be modified are studied), (3) application-specific research, and (4) computing environment-related research (Storn 2008).

The main approach used to improve DE is to combine it with other optimization algorithms (Brest 2009). This mixing of the best features of two or more techniques is called hybridization, the result being a new algorithm which is expected to outperform its parents (components) (Das & Suganthan 2011). Hybridization of the DE algorithm can be performed at four levels: (1) individual, which describes the behavior of an individual; (2) population, which describes the dynamics of a population or subpopulation; (3) external, which describes the interaction with other methods; and (4) meta-level, which includes different algorithms as one of the possible strategies (Feoktistov 2006).

In this paper, in order to improve the performance of the optimizer and to raise the probability of an optimal NN determination, a modified version of DE was hybridized with two algorithms represented by the Random Search and BK. This hybridization is performed at the individual level, in a greedy approach. Based on randomly generated numbers, the best solution found at each generation is improved using one of the two algorithms. Initially, only BK algorithm was used as a local search, but it was observed that, in some cases (for yet unknown reasons), this training algorithm does not

manage to improve the neural network, although it is applied multiple times (for example when the solution does not evolve for a few generations). One way to solve this problem was to change the algorithm applied, but since this problem seldom appears, it was considered, based on practical considerations, that it is better to add a second algorithm for local improvement (Random Search).

The modifications performed to the DE base core are represented by the introduction of the opposition-based learning (OBL) as a mode to improve initialization, by a simple self-adaptive principle which replaced the hand tuning of the control parameters and by the reorganization of the individuals participating in the mutation phase based on the fitness function. These changes, in combination with neural networks, were also used, in different variants, in previous works performed by our group, in connection with various types of applications (Dragoi et al. 2011, 2012a, b, 2013). They were chosen based on the fact that each modification brings a plus of performance on the overall algorithm as it was demonstrated in the abovementioned references. The novelty of the current version consists in combining these modifications with BK and Random Search and applying it to the depollution process of some gaseous compounds.

Since the main idea of the proposed hSADE-NN is to determine the optimum neural network that acts as a model for the chosen process, DE algorithm works with a population of encoded networks. Taking into account that it is an evolutionary algorithm, the population of networks evolves over a number of generations by employing the mutation, crossover, and recombination procedures.

The OBL principle was proposed by Tizhoosh (2005), and it states that the probability that a number is better than its opposite is 50 %. In the current work, a normal distribution is used to generate the potential solution and, after that, its opposite is computed. From this group, the best individuals are chosen to compose the initial population.

The mutation in the classical DE algorithm is represented by the addition to a base vector of a scaled differential term. In literature, various studies related to the use of variants of this principle can be encountered, the results indicating that, depending on the problem being solved, the performance of the mutation changes (Dragoi et al. 2011). Consequently, in the current work, the algorithm is based on a modified approach, using two differential terms and a base vector chosen as the best individual selected from the individuals participating in the mutation phase. From the five distinct mutation individuals, four are randomly chosen and one is represented by the current individual. This approach is used so that all the

networks in the population have an equal chance to participate in the mutation phase.

The crossover type applied here is represented by the binomial version, where the parameters are randomly inherited from one of the two parents. The type of selection is the tournament version, where a one-to-one competition between the individuals from the current and the trial generations takes place. The survival criterion used in the hSADE-NN algorithm is the objective function value only, the networks with the best fit participating in the next generation.

Multiple studies related to the use of DE control parameters reported that the best method to replace the trial and error approach when determining their best values is represented by self-adaptation (Dragoi et al. 2011). In the current work, a simple self-adaptation principle is used. The control parameters are included into the algorithm itself, and they are evolved using the same mathematical equation as the individuals containing them. In this manner, the complexity of the algorithm remains relatively the same as that of the non-adaptive variants.

The type of encoding applied in the current study is represented by direct encoding, where a direct mapping between genotype (the network representation) and phenotype (the actual neural network) exists. A large number of parameters are considered for the optimization, the structure of the encoded networks DE works with being presented in Fig. 2, where $N_h$ represents the number of hidden layers; $N_{h1}$ is the number of neurons in the first hidden layer; $N_{h2}$ is the number of neurons in the second hidden layer; $w_j$, $j=1..\ m$, represents the weights between the neurons; $b_i$, $A_{fi}$, and $P_{afi}$, $i=1\ldots n$ are the biases, activation functions, and parameters of the activation functions, respectively, corresponding to the $i$th neuron of the network.

As it can be observed from Fig. 2, the encoded networks have a relatively high number of parameters which are dependent on the network structure. In order to limit this number, two main aspects are considered: (1) the number of hidden layers is limited to two since it was determined that a two-hidden-layer NN can model, with an acceptable accuracy, almost any process, and (2) the maximum number of neurons allowed in the first and second hidden layers are 40 and 20, respectively. These values were chosen based on practical and empirical considerations, various rules in literature suggesting that the number of neurons in the first hidden layer must be higher than in the second hidden layer. In addition, 40 neurons for the first hidden layer are a maximum limit for the majority of the practical cases. Various simulations performed by our group on multiple chemical engineering processes with a maximum of ten inputs indicated that this interval is sufficient
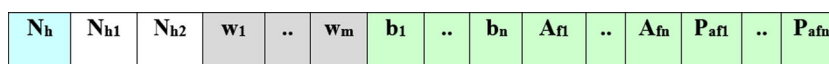
| $N_h$ | $N_{h1}$ | $N_{h2}$ | $w_1$ | .. | $w_m$ | $b_1$ | .. | $b_n$ | $A_{f1}$ | .. | $A_{fn}$ | $P_{af1}$ | .. | $P_{afn}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

**Fig. 2** Structure of the encoded network

for providing good models. In this manner, the number of weights is restricted to $in \times 40 + 40 \times 20 + 20 \times out$ (where $in$ represents the number of inputs and $out$ the number of outputs of the considered neural network). Also, the number of biases, activation functions, and parameters of the activation functions is each limited to a value of $in + 40 + 20 + out$. Consequently, the vector can have $3 + 43 \times in + 980 + 32 \times out$ parameters. For the current case study, where five inputs and one output are considered, the number of vector components is equal to 1,230.

On what concerns the activation functions, five types were applied here: bipolar sigmoid, logarithmic sigmoid, tangent sigmoid, sinus, and triangular basis functions. Various types of activation functions can have distinct effects on the NN performance, their correct choice being a problem for the researchers. Consequently, by introducing them into the genotype, the DE algorithm can efficiently explore the search space and determine the optimal combination of the activation functions which lead to high NN performance. From the selected types of activation functions, only bipolar sigmoid and logarithmic sigmoid require an additional parameter, but, since there is the possibility that each neuron can have one of the five types of activation functions, at the end of the vector, a space is allocated for these values.
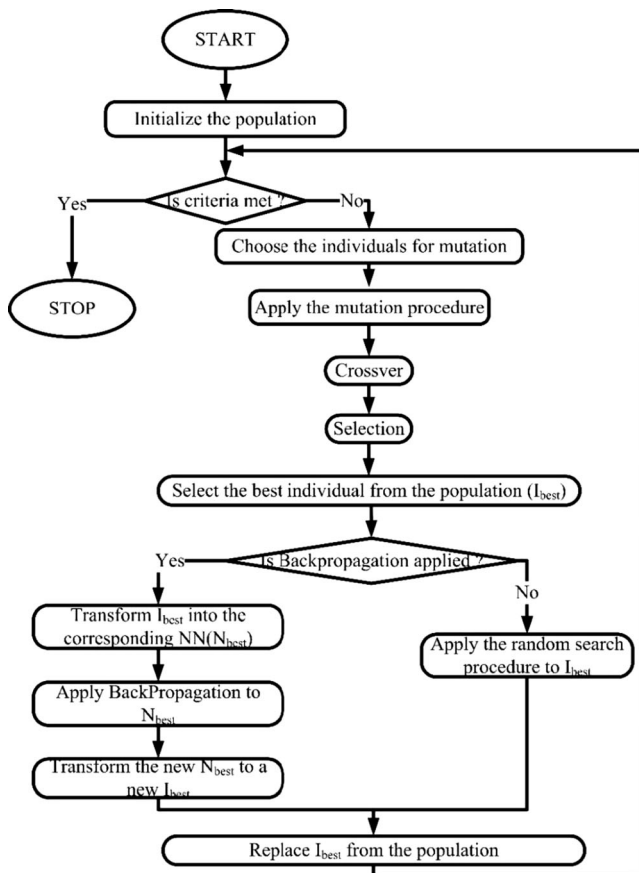


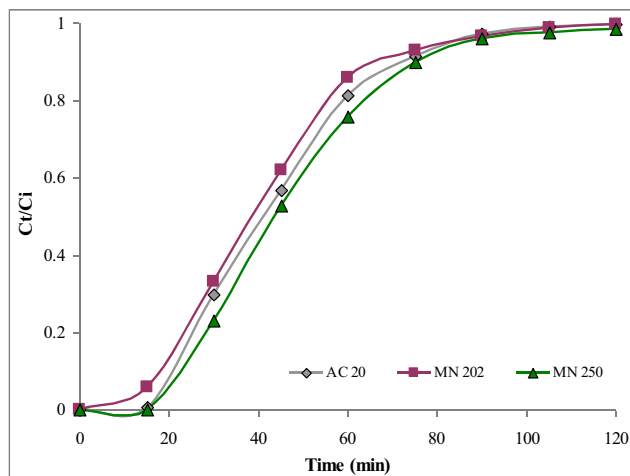**Fig. 3** Schema of the hSADE-NN algorithm



**Fig. 4** The breakthrough curves for *n*-hexane adsorption on all the three adsorbents for the initial VOC concentration of 11 mg/L

Compared with our previous DE-based algorithm, called SADE-NN-2 (Dragoi et al. 2012b), in hSADE-NN, the randomization of the method used for locally improving the best solution represents the main modification performed. On the other hand, SADE-NN-2 was specially designed and applied for monitoring and controlling the freeze-drying process, which has other difficulties than the considered process represented by depollution of some gaseous streams containing volatile organic compounds.

The simplified schema of the hSADE-NN algorithm is presented in Fig. 3. $I_{best}$ represents the best individual in the population, and $N_{best}$ is the neural network corresponding to $I_{best}$. As it can be observed, a series of steps are repeated until a stop criterion is met. In the current work, this criterion is a mixed one and is represented by the number of generations reaching a predefined value and the fitness function reaching a very low value ($10^{-10}$).
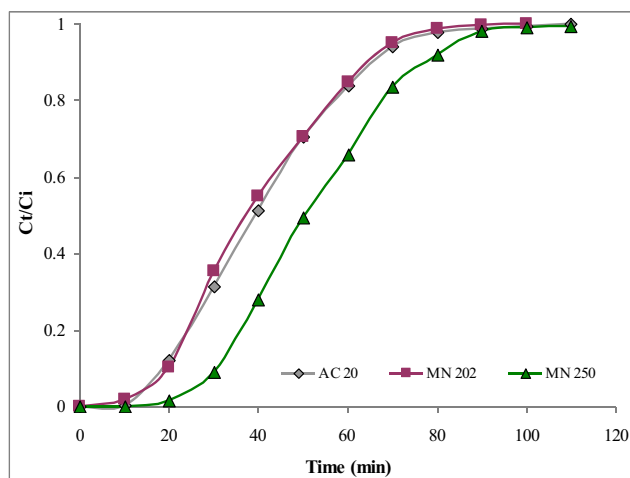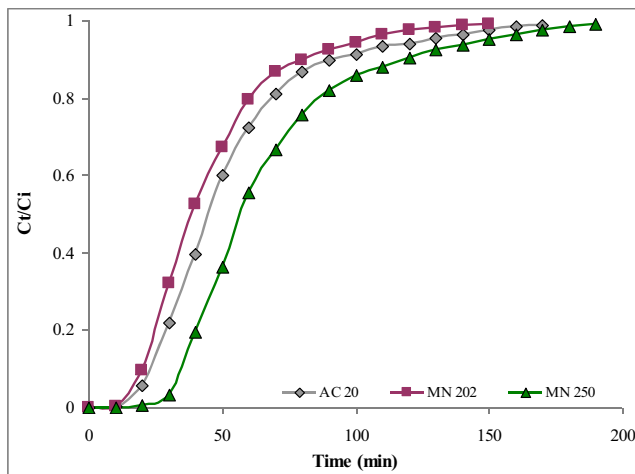


**Fig. 5** The breakthrough curves for *n*-hexane adsorption on all the three adsorbents at 40 °C

**Fig. 6** The breakthrough curves for *n*-hexane adsorption on all the three adsorbents for the adsorbent layer thickness of 1.5 cm

After the classical steps of the DE algorithm (mutation, crossover, and selection) are performed, a new step related to the local search procedure is included. If the random number generated is lower than 0.5, then BK is applied and the selected solution must be transformed from a real number vector into a neural network. If not, the Random Search algorithm is applied, without the need to transform the vector into a network.

The BK algorithm used in this case is the classical variant with no alteration or modifications. In the case of the Random Search procedure, a simple version is considered. Only the parameters corresponding to the weight values are altered using Random Search as they are the most sensitive and small variations can lead to significant improvements. For a small number of iterations, a random value is added to the weight parameters of the initial individual considered for improvement (denoted $x_{ind}$), at the end of each iteration the fitness of the new individual being computed in order to compare it with the one of $x_{ind}$. If it is bigger, then $x_{ind}$ is replaced and the procedure is repeated until the specific number of iterations is reached.

As it can be observed from Fig. 3, BK is applied only after a decoding procedure is performed. This procedure is required due to the fact that the structure of the DE individuals is not compatible with BK, although it contains all the information corresponding to a complete neural network. Therefore, a decoding-encoding procedure is performed each time BK is used as a local search algorithm.

The fitness function used (Eq. 2) for determining the network performance is based on the mean squared error (MSE), and each time it is computed (as in the case of BK application), a transformation DE individual network is required. In this manner, the information flows from DE to neural model and vice-versa.

$$\text{Fitness} = \frac{1}{\text{MSE}_{\text{training}} + \text{err\_correction}} \qquad (2)$$

where err_correction is a constant value (equal to $10^{-10}$) introduced to eliminate the improbable case when $\text{MSE}_{\text{training}}$ is equal to 0.

## Results and discussion

### Processing of experimental data

The experimental data obtained in dynamic conditions were processed to determine the influence of each parameter analyzed, namely temperature, length of the bed, and initial

**Table 1** Results of the simulations performed with hSADE-NN algorithm

|   | Fitness function | Network topology | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|---|
|   |   |   | MSE | PR (%) | r | MSE | PR (%) | r |
| 1 | 16.47858 | 5:18:1 | 0.06068 | 6.53227 | 0.99443 | 0.05814 | 9.22087 | 0.99679 |
| 2 | 16.39714 | 5:15:1 | 0.06098 | 7.12258 | 0.99453 | 0.05694 | 8.75751 | 0.99578 |
| 3 | 16.38152 | 5:15:1 | 0.06104 | 7.11830 | 0.99452 | 0.05718 | 8.77732 | 0.99575 |
| 4 | 16.25329 | 5:17:1 | 0.06152 | 7.08973 | 0.99434 | 0.05847 | 10.84208 | 0.99539 |
| 5 | 16.15935 | 5:26:1 | 0.06188 | 7.05038 | 0.99436 | 0.05946 | 9.41623 | 0.99529 |
| 6 | 16.05451 | 5:19:1 | 0.06228 | 7.01625 | 0.99446 | 0.05581 | 9.85281 | 0.99607 |
| 7 | 16.04701 | 5:29:1 | 0.06231 | 7.78126 | 0.99444 | 0.06607 | 10.45398 | 0.99454 |
| 8 | 16.01581 | 5:20:1 | 0.06243 | 6.40319 | 0.99415 | 0.05498 | 9.34992 | 0.99614 |
| 9 | 15.96990 | 5:16:1 | 0.06261 | 7.34123 | 0.99446 | 0.06009 | 9.78418 | 0.99540 |
| 10 | 15.73425 | 5:18:1 | 0.06356 | 7.50501 | 0.99411 | 0.06152 | 10.27262 | 0.99580 |
| Average | 16.14914 |   | 0.06193 | 7.09602 | 0.99438 | 0.05887 | 9.67275 | 0.99569 |

**Table 2** Results obtained with SADE-NN-2 algorithm

|  | Fitness function | Network topology | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|---|
|  |  |  | MSE | PR (%) | r | MSE | PR (%) | r |
| 1 | 15.69654 | 5:13:1 | 0.06374 | 8.15766 | 0.99419 | 0.06036 | 10.44920 | 0.99524 |
| 2 | 15.55176 | 5:20:1 | 0.06430 | 8.79539 | 0.99380 | 0.06340 | 12.11586 | 0.99487 |
| 3 | 15.51521 | 5:13:1 | 0.06445 | 8.26029 | 0.99391 | 0.05912 | 11.34787 | 0.99546 |
| 4 | 15.49429 | 5:09:1 | 0.06454 | 9.15480 | 0.99408 | 0.06759 | 12.93116 | 0.99434 |
| 5 | 15.43640 | 5:20:1 | 0.06478 | 9.18417 | 0.99160 | 0.06617 | 13.00828 | 0.99240 |
| 6 | 15.33790 | 5:15:1 | 0.06520 | 8.69051 | 0.99407 | 0.06374 | 10.97239 | 0.99469 |
| 7 | 15.27924 | 5:09:1 | 0.06544 | 8.39316 | 0.99407 | 0.06245 | 10.98566 | 0.99497 |
| 8 | 15.23317 | 5:39:1 | 0.06564 | 8.77666 | 0.99387 | 0.06852 | 11.64928 | 0.99388 |
| 9 | 15.03166 | 5:17:1 | 0.06652 | 8.31172 | 0.99356 | 0.06591 | 12.02195 | 0.99460 |
| 10 | 14.94833 | 5:12:1 | 0.06759 | 9.34549 | 0.09935 | 0.06658 | 12.39459 | 0.99452 |
| Average | 15.35245 |  | 0.06522 | 8.70699 | 0.90425 | 0.06438 | 11.78762 | 0.99450 |

concentration. The amount of pollutant retained increases with increasing concentration for all the three adsorbents studied. All adsorbents have a good behavior at low concentrations, respectively at 7 mg/L, when the ability to retain pollutant is higher, the breakthrough curves follow the "S" shape and the breakpoint is reached later.

Comparing the amount of $n$-hexane adsorbed at a time, i.e., the same initial concentration of 11 mg/L, good behavior has the hypercross-linked polymer MN 250 (Fig. 4). $C_i$ is the inlet concentration of $n$-hexane (ppmv), and $C_t$ is the outlet concentration of $n$-hexane at arbitrary time (ppmv).

With increasing temperature, the adsorption capacity of MN 202, MN 250, and AC 20 for $n$-hexane vapors decreases and it was observed that it is most pronounced for the MN 250 adsorbent. The decrease of adsorption capacity with the temperature indicates the exothermic nature of the adsorption of $n$-hexane vapors on these adsorbents. Also, with increasing temperature, the breakthrough curves became steeper and the breakpoint time is reached earlier for each adsorbent

considered. It was also observed that the great performance for $n$-hexane vapor adsorption with increasing temperature was recorded by the MN 250 adsorbent, as Fig. 5 shows for a temperature of 40 °C.

The experiments also demonstrate that the length increase of the adsorbent bed provides a better adsorption percentage and a higher breakpoint time for each adsorbent studied. Concerning the type of material used as adsorbent, a behavior similar to that of previous cases can be observed, when the best performance for the adsorption of $n$-hexane vapors is recorded by MN 250 hypercross-linked polymeric resin, followed by granular activated carbon AC 20, and then by hypercross-linked polymeric resin MN 202 (Fig. 6).

Neural network modelling

After the experimental data were gathered and processed using the normalization procedure and practical considerations related to the training/testing neural networks were

**Fig. 7** Evolution of $MSE_{training}$ and $MSE_{testing}$ during the determination of a neural network model using hSADE-NN algorithm
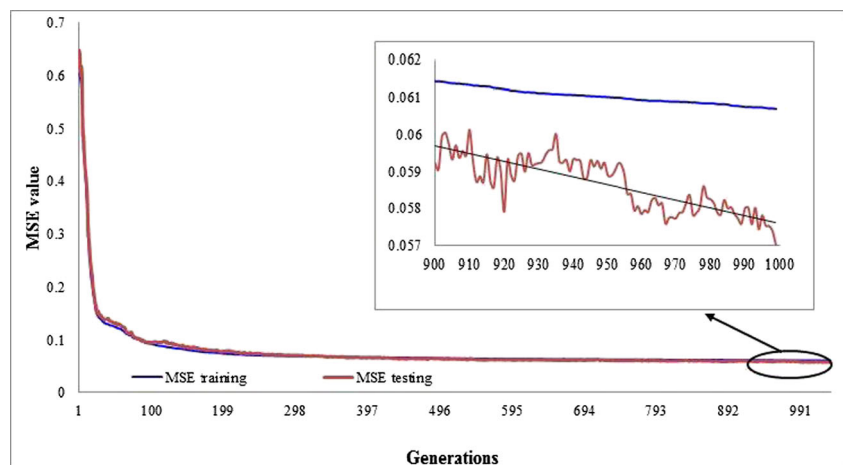
**Fig. 8** Evolution of $MSE_{training}$ and $MSE_{testing}$ during the determination of a neural network model using SADE-NN-2 algorithm
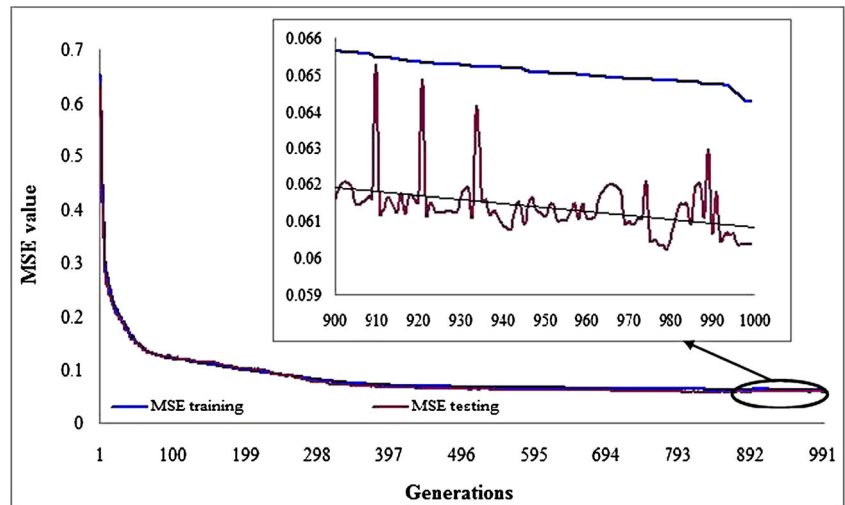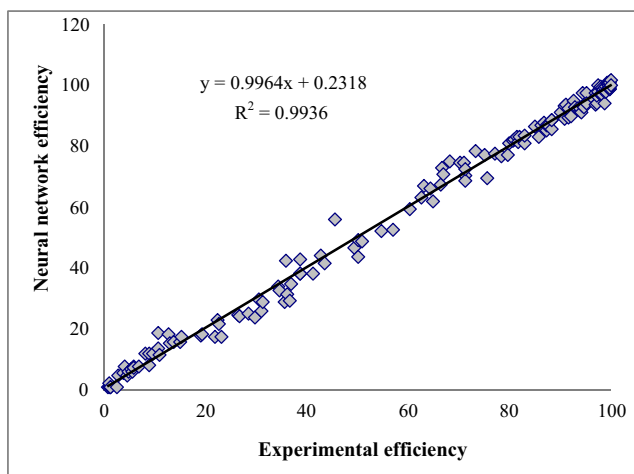


applied, the next step was represented by the determination of an optimal model for the process. This task was achieved by performing various software simulations using the hSADE-NN algorithm proposed in this work.

As previously mentioned, a series of limitations to the network structure were imposed (maximum of two hidden layers with 40 and 20 neurons, respectively) to reduce the individual complexity.

A series of ten results ordered by the fitness function are presented in Table 1, were the network topology is coded using the following structure: in:$N_{h1}$:$N_{h2}$:out, MSE is the mean square error, $r$ is the correlation coefficient between experimental and simulation data, and PR is the average percentage error.

The average percentage error was computed using the following formula:

$$PR = \frac{100}{n} \sum_{1}^{n} \frac{x_{i,0} - x_i}{x_i} \quad (3)$$

where $x_i$ is the predicted value and $x_{i,\,0}$ is the $i$th measured value.

As it can be observed from Table 1, although the network settings allow the determination of two-hidden-layer neural networks, the considered algorithm found only one-hidden-layer networks. This indicates that the proposed algorithm determines models with a simple structure. On the other hand, an average value of MSE around 0.06 in both training and testing phases represents an acceptable performance of the model.

By comparing the data from Table 1 with the ones from Table 2 (results obtained with previous SADE-NN-2), it can be observed that, from various points of view (fitness value, MSE, average relative error, correlation), hSADE-NN is better in finding the suitable neural network model. On what concerns the network topology, SADE-NN-2 seems to determine simpler models, but since their performance is not as good as the one determined with hSADE-NN (which has only a few more neurons), the former algorithm is considered providing the best results. In this way, the creation of a new, improved version of the DE algorithm is justified.

The best networks are the ones having the highest fitness: value of 16.478 and network 5:18:1 generated by hSADE-NN and 15.696 for network 5:13:1, generated by SADE-NN-2. For determining if these best networks are overtrained, a comparison between the evolution of $MSE_{training}$ and $MSE_{testing}$ during simulations was realized (Fig. 7 for hSADE-NN and Fig. 8 for SADE-NN-2). This action is performed for safety, although the possibility of overtraining



**Fig. 9** MLP(5:18:1) network prediction vs. experimental data in the testing phase

**Table 3** Yoon-Nelson parameters for MN 202, MN 250, and AC 20 adsorbents (Buburuzan et al. 2010)

| Adsorbent | $k_{YN}$ (min$^{-1}$) | $\tau$ (min) |
|---|---|---|
| MN 202 | 0.1393 | 27.4064 |
| MN 250 | 0.1242 | 39.583 |
| AC 20 | 0.1367 | 34.1902 |

**Fig. 10 a** NN predictions compared with experimental data and **b** Yoon-Nelson model results compared with experimental data, for $T=30$ °C and a bed length of 1 cm

$y = 0.9999x - 0.5$
$R^2 = 0.9951$

$y = 1.0124x - 2.1311$
$R^2 = 0.9854$

(a)

(b)
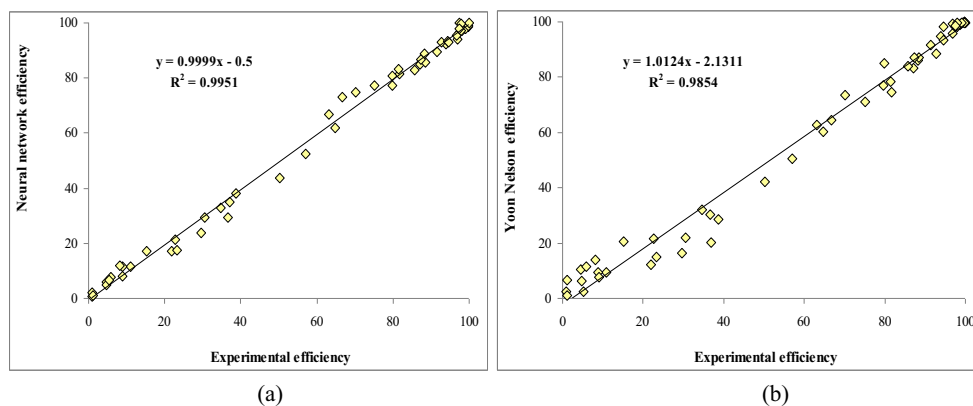
a neural network when using hSADE-NN algorithm or its predecessors is quite low due to the following factors: (1) a simultaneous topological and structural optimization is performed and (2) the population of networks evolves, the best solution not always representing the same individual over the entire generations.

Comparing Figs. 7 and 8, it can be observed that the curves of MSE evolution in both training and testing phases are similar in the first 20 % of the entire number of generations, registering the most abrupt change. In the other 80 %, the networks evolved slower, the individuals being placed in local optimum from which the algorithm is trying to escape.

One can see in Figs. 7 and 8 that after 200 generations, the algorithm can stop since there is no significant improvement rate. On the other hand, since the desired outcome is a near-optimal neural network which models the considered process, the number of simulations performed is 1,000. As the current generation grows, the $MSE_{training}$ becomes smaller and smaller, but the $MSE_{testing}$ varies, although it follows a descending trend as indicated by the trend lines from Figs. 7 and 8. This evolution shows that the determined network is not overtrained, a good correlation between the desired data and the one generated with the 15:18:1 network (the results of hSADE-NN) existing in the testing phase (Fig. 9).

Concerning the MSE in the testing phase of the two algorithms, for the duration of the last 10 % of evolution, hSADE-NN has a steeper trend line, indicating that if the evolution is extended (by raising the maximum allowed number of generations), then a better network will be found. Since this improvement is small and the computational resources consumed rise, a compromise is considered, the current best network MLP(5:18:1) being chosen as a suitable model for the process of depollution of gaseous streams containing volatile organic compounds.

This steeper trend line observed for hSADE-NN is the result of a second local search algorithm. By alternating two different local algorithms, the probability of creating an individual which can introduce a new direction of search in the mutation phase is higher. Consequently, the search space is

better scanned for feasible areas where the optimum can be found.

On what concerns the fact that our group developed a series of DE versions (Dragoi et al. 2011, 2012a, b, 2013), it is worth mentioning that, depending on the process on which they are applied, the DE variants could have different performances. Consequently, there is a possibility of choosing, testing, and selecting the best variant, according to the accuracy of the results.

## Comparison between NN and classical models

The model developed by Yoon and Nelson (1984), which is a descriptive model that uses experimental data to calculate its parameters, was applied to investigate the breakthrough behavior of $n$-hexane vapors onto the three adsorbents, AC 20, MN 202, and MN 250. The model has been developed based on the assumption that the decrease of the probability of adsorption for each adsorbate molecule is proportional to the probability of adsorbate adsorption and the probability of adsorbate breakthrough on the adsorbent (Hu et al. 2009;
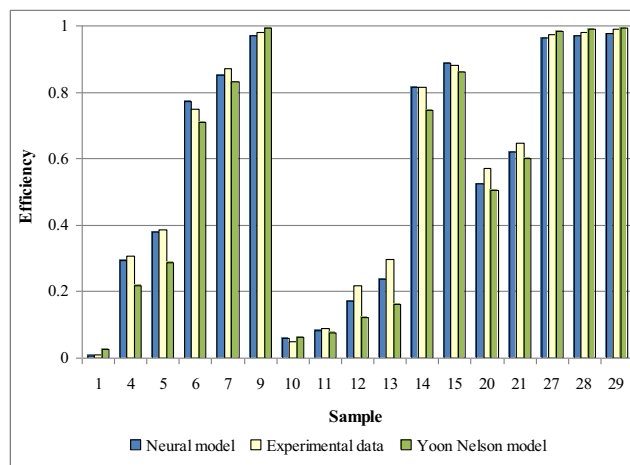
**Fig. 11** Comparison between the experimental data set, Yoon-Nelson model, and neural network predictions

Ahmad & Hameed 2010). The equation describing the model is

$$\ln \frac{C_1}{C_1 - C_1} = k_{YN} \times t - \tau \times k_{YN} \qquad (4)$$

where $k_{YN}$ is the rate velocity constant ($min^{-1}$), $\tau$ is the time required for 50 % adsorbate breakthrough (min), and $t$ is the breakthrough (sampling) time (min).

The values $k_{YN}$ and $\tau$ were determined from $\ln[C_t/(C_i-C_t)]$ against $t$ plots, at a flow rate of 130 $cm^3$/min, a bed length of 1 cm for each adsorbent, and a space velocity of 2.75 $s^{-1}$. These values are listed in Table 3 for MN 202, MN 250, and AC 20 adsorbents.

Figure 10 presents the results provided by MLP(5:18:1) (Fig. 10a) and Yoon-Nelson model (Fig. 10b) against experimental data. Only an example is presented here ($T$=30 °C and bed length=1 cm), but the conclusion is the same for the whole set of data.

The mean square error was calculated for each situation: MSE=0.002565 for NN modelling, where NN is determined with hSADE-NN, and MSE=0.007680 for the Yoon-Nelson model (three times higher). The results are comparable in average values, but the use of a descriptive model is more laborious and has to be adapted to each set of experimental data. Once developed and trained, the NN model is able to make predictions in a facile manner and in a reasonable time interval. In addition, the NN model can be easily integrated in on-line optimization procedures, being capable to work with different optimization algorithms.

The benefits of NN modelling are more evident in Fig. 11 where the comparison is made point by point. The fact that the predictions of MLP(5:18:1) are closer to experimental data than the results provided by the Yoon-Nelson model is visible. A similar conclusion derives from Fig. 10—the phenomenological model results are more scattered from the trend line (Fig. 10b). Consequently, the neural model can be considered an efficient modelling alternative, useful for experimental practice.

## Conclusions

In this research, three different adsorbents were studied for adsorption of $n$-hexane vapors from the gaseous stream, in the dynamic regime—the granular activated carbon AC 20 and two hypercross-linked polymeric resins, MN 202 and MN250. The experimental study was carried out by varying some important parameters that influence the efficiency of the adsorption process like concentration of VOCs in the gaseous stream, temperature, and length of the adsorbent bed in the adsorption column. Generally, the three adsorbents recorded

high efficiency for $n$-hexane vapors retention from gaseous stream, the performance of adsorption decreasing according to the hierarchy: MN 250>AC 20>MN 202.

The depollution process of gaseous streams is a complex problem which poses a series of problems related to the determination of a good model. In the current work, this problem is solved using a neuro-evolutionary approach that combines a differential evolution algorithm with a local search procedure employing two distinct local algorithms: Random Search and Backpropagation. These algorithms are randomly used to improve the best solution generated at each generation by DE. The DE algorithm included an opposition-based learning initialization, a simple self-adaptive method for controlling the parameter tuning, and a modified mutation principle based on the fitness function as a criterion for reorganization. This variant of DE combined with local search algorithms leads to the new hSADE-NN which, as the results indicated, is able to find good neural network models for the considered process, in an optimum variant.

The best neural model was also compared with a classical model, the results indicating that the neural network predictions are closer to the experimental data. Thus, the proposed algorithm is a viable alternative to the classical approaches used in modelling of the depollution process of gaseous streams.

hSADE-NN can be also considered an efficient and flexible algorithm, having the capability to be applied successfully to other different processes and systems, as it was designed in a general form.

## References

Ahmad AA, Hameed BH (2010) Fixed-bed adsorption of reactive azo dye onto granular activated carbon prepared from waste. J Hazard Mater 175:298–303

Brest, J. (2009). Constrained real-parameter optimization with e-self-adaptive differential evolution. In E. Mezura-Montes (Ed.), Constraint-handling in evolutionary optimization. Studies in Computational Intelligence Springer Berlin/Heidelberg, pp. 73-93.

Buburuzan AM, Catrinescu C, Macoveanu M (2009) Adsorption of n-hexane vapors onto non-functionalised hypercrosslinked polymers (hypersol-macronettm) and activated carbon: thermodynamic and kinetic studies. Environ Eng Manag J 8:259–265

Buburuzan AM, Catrinescu C, Macoveanu M (2010) Comparative study of the adsorption-desorption cycles of hexane over hypercrosslinked polymeric adsorbents and activated carbon. Environ Eng Manag J 9: 125–132

Caicedo Torres W, Quintana M, Pinzon H (2013) Differential diagnosis of hemorrhagic fevers using ARTMAP and artificial immune system. Int J Artif Intell 11:150–169

Chandra A, Yao X (2006) Ensemble learning using multi-objective evolutionary algorithms. J Math Model Algoritm 5:417–445

Curteanu S, Cartwright H (2012) Neural networks applied in chemistry. I Determination of the optimal topology of multilayer perceptron neural networks J Chemom 25:527–549

Das S, Suganthan PN (2011) Differential evolution: a survey of the state-of-the-art. IEEE Trans Evol Comput 15:4–31

Dragoi EN, Curteanu S, Leon F, Galaction AI, Cascaval D (2011) Modeling of oxygen mass transfer in the presence of oxygen-vectors using neural networks developed by differential evolution algorithm. Eng Appl Artif Intell 24:1214–1226

Dragoi EN, Curteanu S, Lisa C (2012a) A neuro-evolutive technique applied for predicting the liquid crystalline property of some organic compounds. Eng Optim 44:1261–1277

Dragoi EN, Curteanu S, Fissore D (2012b) Freeze-drying modeling and monitoring using a new neuro-evolutive technique. Chem Eng Sci 72:195–204

Dragoi EN, Curteanu S, Galaction AI, Cascaval D (2013) Optimization methodology based on neural networks and self-adaptive differential evolution algorithm applied to an aerobic fermentation process. Appl Soft Comput 13:222–238

Feoktistov, V. (2006). Differential evolution: in search of solutions. Springer

Floreano D, Durr P, Mattiussi C (2008) Neuroevolution: from architectures to learning. Evol Intell 1:47–62

Furtuna R, Curteanu S, Leon F (2012) Multi-objective optimization of a stacked neural network using an evolutionary hyper-heuristic. Appl Soft Comput 12:133–144

Gupta VK, Verma N (2002) Removal of volatile organic compounds by cryogenic condensation followed by adsorption. Chem Eng Sci 57:2679–2696

Halteta Buburuzan MB, Catrinescu C, Macoveanu M (2009) Adsorption of n-hexane vapors onto non-functionalized hypercrosslinked polymers (hypersol-macronettm) and activated carbon: equilibrium studies. Environ Eng Manag J 8:173–181

Han L, Shi X, Wu W, Kirk FL, Luo J, Wang L et al (2005) Nanoparticle-structured sensing array materials and pattern recognition for VOC detection. Sens Actuators B: Chem 106:431–441

Hernandez RP, Alvarez-Gallegos J, Reyes J (1998) Simple recurrent neural network: a neural network structure for control systems. Neurocomput 3:277–289

Hu Q, Li JJ, Hao ZP, Li LD, Qiao SZ (2009) Dynamic adsorption of volatile organic compounds on organofunctionalized SBA-15 materials. Chem Eng J 149:281–288

Khan FI, Ghoshal A (2000) Removal of volatile organic compounds from polluted air. J Loss Prev Process Ind 13(6):527–545

Kim KJ, Kang CS, You YJ, Chung MC, Woo MW, Jeong WJ et al (2006) Adsorption-desorption characteristics of VOCs over impregnated activated carbons. Catal Today 111:223–228

Leeghim H, Seo IH, Bang H (2008) Adaptive nonlinear control using input normalized neural networks. J Mech Sci Technol 22:1073–1083

Lillo-Rodenas MA, Fletcher AJ, Thomas KM, Cazorla-Amoris D, Linares-Solano A (2006) Competitive adsorption of a benzene-toluene mixture on activated carbons at low concentration. Carbon 44:1455–1463

Llanos J, Rodrigo MA, Canizares P, Furtuna RP, Curteanu S (2013) Neuro-evolutionary modelling of the electrodeposition stage of a polymer-supported ultrafiltration-electrodeposition process for the recovery of heavy metals. Environ Model Softw 42:133–142

Matros YS, Noskov AS, Chumachenko VA, Goldman OV (1988) Theory and application of unsteady catalytic detoxication of effluent gases from sulfur dioxide, nitrogen oxides and organic compounds. Chem Eng Sci 43:2061–2066

Matthies M, Giupponi C, Ostendorf B (2007) Environmental decision support systems: current issues, methods and tools. Environ Model Softw 22:123–127

Neri F, Tirronen V (2010) Recent advances in differential evolution: a survey and experimental analysis. Artif Intell Rev 33:61–106

Odabasi M, Ongan O, Cetin E (2005) Quantitative analysis of volatile organic compounds (VOCs) in atmospheric particles. Atmospheric Environ 39:3763–3770

Pant M, Thangaraj R, Abraham A, Grosan C, Differential Evolution with Laplace mutation operator (2009) Proceedings of the Eleventh Conference on Congress on Evolutionary Computation, Trondheim. IEEE Press, Norway, pp 2841–2849

Patan K, Approximation Abilities of Locally Recurrent Networks (2008) Artificial neural networks for the modelling and fault diagnosis of technical processes. Lecture notes in control and information sciences. Springer Berlin, Heidelberg, pp 65–75

Peng L, Wang Y (2010) Differential evolution using uniform-quasi-opposition for initializing the population. Inf Technol J 9:1629–1634

Pirdashti M, Curteanu S, Kamangar MH, Hassim M, Amid M (2013) Artificial neural networks: applications in chemical engineering. Rev Chem Eng 29:205–239

Precup RE, Tomescu ML, Preitl S (2007) Lorenz system stabilization using fuzzy controllers. Int J Comput Commun Control 2:279–287

Shim WG, Lee JW, Moon H (2006) Adsorption equilibrium and column dynamics of VOCs on MCM-48 depending on pelletizing pressure. MicroporousMesoporous Mater 88:112–125

Silvestre-Albero A, Ramos-Fernandez JM, Martinez-Escandell M, Seplveda-Escribano A, Silvestre-Albero J, Rodriguez-Reinoso F (2010) High saturation capacity of activated carbons prepared from mesophase pitch in the removal of volatile organic compounds. Carbon 48:548–556

Srivastava AK (2003) Detection of volatile organic compounds (VOCs) using SnO2 gas-sensor array and artificial neural network. Sensors Actuators B Chem 96:24–37

Storn R, Differential Evolution Research - Trends and Open Questions. In U. Chakraborty (2008) Advances in differential evolution. Studies in computational intelligence. Springer Berlin, Heidelberg, pp 1–31

Subudhi B, Jena D (2009) An improved differential evolution trained neural network scheme for nonlinear system identification. Int J Autom Comput 6:137–144

Tchoupo GN, Guiseppi-Elie A (2005) On pattern recognition dependency of desorption heat, activation energy, and temperature of polymer-based VOC sensors for the electronic NOSE. Sensors Actuators B Chem 110:81–88

Tizhoosh HR (2005) Opposition-based learning: a new scheme for machine intelligence. In: In: International Conference on Computational Intelligence for Modeling, Control and International Conference on Intelligent Agents, Web technologies and Internet Commerce. 28-30 Nov 2005., pp 695–701

Vassileiou A, Maris F, Kitikidou K, Iliadis L (2012) Artificial neural networks for improved predictions in flow estimation. Int J Artif Intell 9:186–201

Xin Y (1999) Evolving artificial neural networks. Proc IEEE 87:1423–1447

Yoon YH, Nelson JH (1984) Application of gas adsorption kinetics. I. A theoretical model for respirator cartridge service life. Am Ind Hyg Assoc J 45:509–516

Zarth, A., & Ludermir, T. B. (2009). Optimization of neural networks weights and architecture: a multimodal methodology. In: Ninth International Conference on Intelligent Systems Design and Applications (ISDA'09) pp. 209-214.